Assessing the Impacts of Environmental Stressors on Macroinvertebrate Indicators in Ohio

Sarada Majumder[1], Florence Fulk[2], Susan Cormier[2] ; [1]SoBran Environmental, c/o [2]National Exposure Research Laboratory, USEPA, 26 W. M.L. King Drive Cincinnati, OH 45268

Abstract

The purpose of this study was to explain and predict the impacts of environmental stressors on macroinvertebrate water quality indicators in the state of Ohio. The ICI and its component metrics developed by the Ohio EPA were used along with environmental quality data for the state of Ohio. As a first step, the impact of each stressor on the ICI and each component metric was identified. At the second step, statistical models were developed to predict the probability of criterion attainment for ICI and its components as a function of the environmental stressors and the interactions between them. The statistical modeling was done using a variant of the logistic regression procedure. The results indicate that in most cases biocriterion attainment can be predicted correctly at 60-65% of the sites using the environmental quality indicators. As expected, uncertainty of prediction increases at extreme values of the environmental stressors. These results can potentially be used for identification of sites with a high risk of water quality degradation in terms of macroinvertebrate indicators, and for formal or informal "threshold" development of the stressors.

Objectives

- Understand the effects of environmental stressors on macroinvertebrates
- Develop models to describe the impact of the stressors on the biotic index (ICI) and its component metrics

Invertebrate Community Index - ICI

- ICI developed by Ohio EPA consists of ten metrics that are sensitive in varying degrees to different categories of impacts
- Metric scores are assigned to each site after calibration for regional differences
- Sites scored 6, 4, 2, 0 for each metric indicating best to worst quality
- ICI is the sum of the metric scores and can range between 0 and 60. Scores that represent biocriterion attainment for Ohio are

Excellent	> 48
Good	32 - 46
Fair	14 – 30
Poor	2 – 12

Macroinvertebrate sampling locations used for this study

Methodology

- Data were compiled representing different categories of stressors, water chemistry, stream habitat quality (Qualitative Habitat Evaluation Index, QHEI, and metrics), riparian land use characteristics, and point source effluents. Analysis was conducted in two steps

Variables used in study

ICI and component metrics Source: Ohio EPA	Water Chemistry Source: STORET	Stream habitat quality Source: Ohio EPA	Riparian characteristics Source: USGS	Point source effluents Source : LEAPS
Total number of Taxa	Ammonia	QHEI	Residential	Ammonia
Number mayfly taxa	BOD	Substrate	Total developed	BOD
Number caddisfly taxa	COD	Cover	Forest	COD
Number dipteran taxa	Total suspended solids	Channel		Fecal coliform
Percent mayfly	Total Kjeldahl nitrogen	Riparian		Phosphorus
Percent caddisfly	pH	Pool		
Percent tanytarsini	Phosphorus	Riffle		
Percent other dipteran	Iron	Gradient		
Percent tolerant	Sodium	Drainage Area		
Qualitative EPT taxa	Magnesium			
ICI attainment status				

Step 1

- Are there any differences between the means of the various stressors among the categories of ICI and its component metrics?

To answer this question, one way ANOVA was used. Since there were a large number of sites with data for certain chemistry variables and a smaller number for others, "effect sizes" were used rather than statistical significance to ascertain impact. This is an absolute measure of the dispersion of the group means compared to the standard deviation of the population. Higher numbers indicate a larger proportion of the population variance for each stressor being accounted for by their membership in the ICI metric groups. For example, 71.2% of the total variance in ammonia concentration is being accounted for by the categories of the "total number of taxa" metric of ICI. Signs added to show direction of relationship.

Step 1 : Results

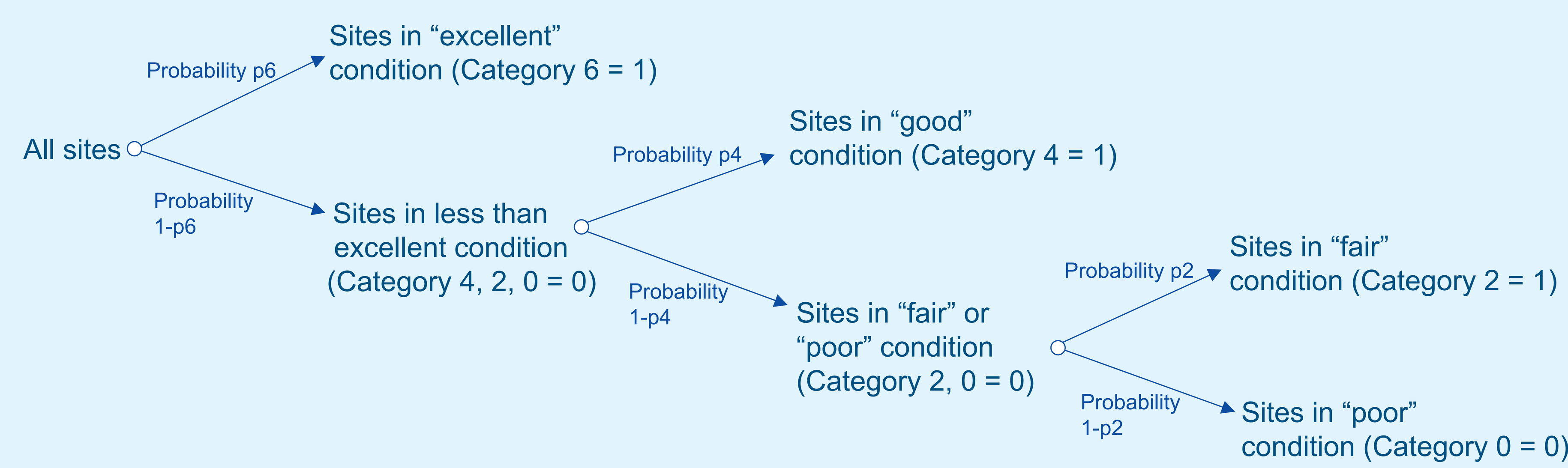
ICI and component metrics : Sensitivity to stressors calculated as proportion of total stressor variance accounted for by the score categories

ICI metrics	Water Chemistry variables										Stream Habitat Variables									Land use variables		
	NH3	BOD	COD	TSS	TKN	pH	Total P	Fe	Na	Mg	QHEI	Substrate	Cover	Channel	Riparian	Pool	Riffle	Gradient	Drainage Area	Residential	Developed land (Urban)	All forest
Total number of Taxa	(-) 0.712	(-) 0.113	(-) 0.006	(+) 0.002	(-) 0.331	(+) 0.333	(-) 0.379	0.000	(-) 0.001	(-) 0.001	(+) 0.384	(+) 0.331	(+) 0.216	(+) 0.331	(+) 0.126	(+) 0.293	(+) 0.339	(+) 0.167	(+) 0.375	(-) 0.118	(-) 0.152	(+) 0.112
Number of mayfly taxa	(-) 0.688	(-) 0.104	(-) 0.011	(+) 0.002	(-) 0.346	(+) 0.357	(-) 0.434	0.000	(-) 0.001	(-) 0.044	(+) 0.387	(+) 0.349	(+) 0.206	(+) 0.338	(+) 0.159	(+) 0.288	(+) 0.363	(+) 0.201	(+) 0.275	(-) 0.152	(-) 0.196	(+) 0.181
Number of caddisfly taxa	(-) 0.493	(-) 0.083	(-) 0.006	(+) 0.002	(-) 0.087	(+) 0.254	(-) 0.34	0.000	(-) 0.000	(-) 0.001	(+) 0.357	(+) 0.258	(+) 0.188	(+) 0.336	(+) 0.198	(+) 0.22	(+) 0.297	(+) 0.239	(+) 0.269	(-) 0.104	(-) 0.159	(+) 0.118
Number of dipteran taxa	(-) 0.524	(-) 0.088	(-) 0.005	(+) 0.000	(-) 0.248	(+) 0.233	(-) 0.281	0.000	(-) 0.001	(-) 0.004	(+) 0.387	(+) 0.326	(+) 0.244	(+) 0.397	(+) 0.131	(+) 0.259	(+) 0.274	(+) 0.21	(+) 0.27	(-) 0.058	(-) 0.067	(+) 0.124
Percent mayfly composition	(-) 0.663	(-) 0.096	(-) 0.006	(+) 0.004	(-) 0.294	(+) 0.342	(-) 0.305	0.000	(-) 0.002	(+) 0.004	(+) 0.271	(+) 0.234	(+) 0.175	(+) 0.265	(+) 0.047	(+) 0.183	(+) 0.257	(+) 0.054	(+) 0.255	(-) 0.123	(-) 0.146	(+) 0.133
Percent caddisfly composition	(-) 0.378	(-) 0.075	(-) 0.005	(+) 0.006	(-) 0.175	(+) 0.128	(-) 0.264	0.000	(-) 0.001	(-) 0.002	(+) 0.325	(+) 0.226	(+) 0.248	(+) 0.29	(+) 0.16	(+) 0.229	(+) 0.312	(+) 0.172	(+) 0.136	(-) 0.105	(-) 0.148	(+) 0.137
Percent tanytarsini	(-) 0.468	(-) 0.098	(-) 0.01	(-) 0.007	(-) 0.322	(+) 0.108	(-) 0.522	0.000	(-) 0.000	(-) 0.004	(+) 0.33	(+) 0.235	(+) 0.186	(+) 0.26	(+) 0.128	(+) 0.229	(+) 0.354	(+) 0.234	(+) 0.109	(-) 0.136	(-) 0.196	(+) 0.197
Percent other dipteran	(-) 0.508	(-) 0.098	(-) 0.007	(+) 0.009	(-) 0.273	(+) 0.181	(-) 0.405	0.000	(-) 0.002	(+) 0.002	(+) 0.362	(+) 0.298	(+) 0.239	(+) 0.27	(+) 0.159	(+) 0.238	(+) 0.416	(+) 0.103	(+) 0.247	(-) 0.158	(-) 0.236	(+) 0.168
Percent tolerant organisms	(-) 0.611	(-) 0.167	(-) 0.008	(+) 0.003	(-) 0.274	(+) 0.246	(-) 0.448	0.000	(-) 0.001	(+) 0.003	(+) 0.417	(+) 0.337	(+) 0.281	(+) 0.365	(+) 0.194	(+) 0.256	(+) 0.387	(+) 0.139	(+) 0.167	(-) 0.208	(-) 0.293	(+) 0.18
Qualitative EPT taxa	(-) 0.763	(-) 0.104	(-) 0.011	(-) 0.005	(-) 0.39	(+) 0.341	(-) 0.506	0.000	(-) 0.002	(-) 0.003	(+) 0.45	(+) 0.448	(+) 0.142	(+) 0.382	(+) 0.257	(+) 0.204	(+) 0.459	(+) 0.297	(+) 0.195	(-) 0.141	(-) 0.228	(+) 0.164
ICI attainment category	(-) 0.79	(-) 0.119	(-) 0.009	(+) 0.003	(-) 0.382	(+) 0.355	(-) 0.505	0.000	(-) 0.002	(+) 0.002	(+) 0.397	(+) 0.253	(+) 0.295	(+) 0.394	(+) 0.287	(+) 0.347	(+) 0.375	(+) 0.148	(+) 0.238	(-) 0.053	(-) 0.082	(+) 0.155

Step 2

- Can we predict the probability of biocriterion attainment using the environmental impact information?

To model ICI and metric responses (e.g. 6, 4, 2, 0), a variant of logistic regression was developed. Logistic regression is used to model the probability of two mutually exclusive events occurring at each site, usually the events are coded 1 if they are desirable and 0 if undesirable. For four categories, three equations are obtained that each predict the probability of falling in category 6 or less than 6, 4 or less than 4, and 2 or 0. The nested coding scheme for the sites and examples of results are shown below.



Step 2 : Results

- Example of results for ICI attainment criterion.

$\log(p6/(1-p6)) = -7.17 + 0.13(\text{Channel}) - 0.88(\log(\text{TKN})) - 0.171(\text{PH}*\log(\text{NH3})) - 0.03(\text{Developed land})$

$\log(p4/(1-p4)) = -20.51 + 2.13(\text{pH}) - 0.77(\log(\text{NH3})) + 0.28(\text{Riffle}) + 0.25(\text{Slope})$

$\log(p2/(1-p2)) = -2.95 - 0.89(\log(\text{TKN})) + 0.2(\text{Channel}) + 0.031(\text{Forest})$

Total chi square = 112, df = 11
Forward stepwise variable selection
Overall predictive efficiency = 65% compared to 25% using random assignment

- Example of results for "number mayfly score" categories

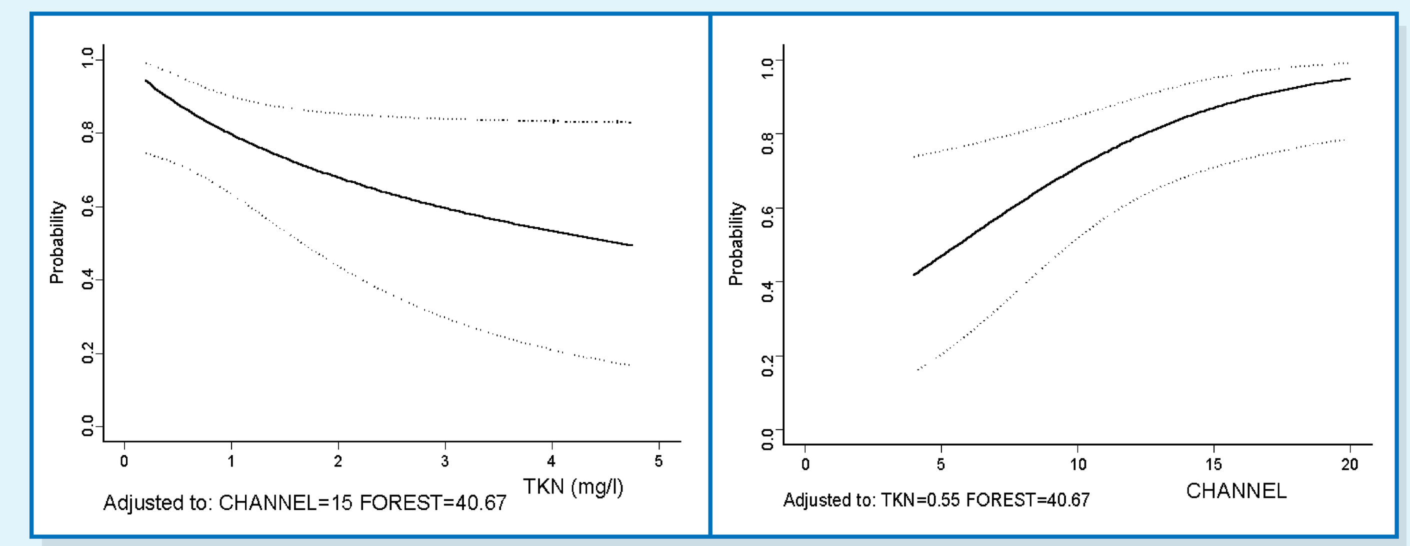
$\log(p6/(1-p6)) = -6.175 + 0.281(\text{Riffle}) - 0.815(\log(\text{TKN})) - 0.189(\text{PH}*\log(\text{NH3})) - 0.02(\text{Developed land})$

$\log(p4/(1-p4)) = -5.14 - 0.065(\text{PH}*\log(\text{NH3})) + 0.216(\text{Channel})$

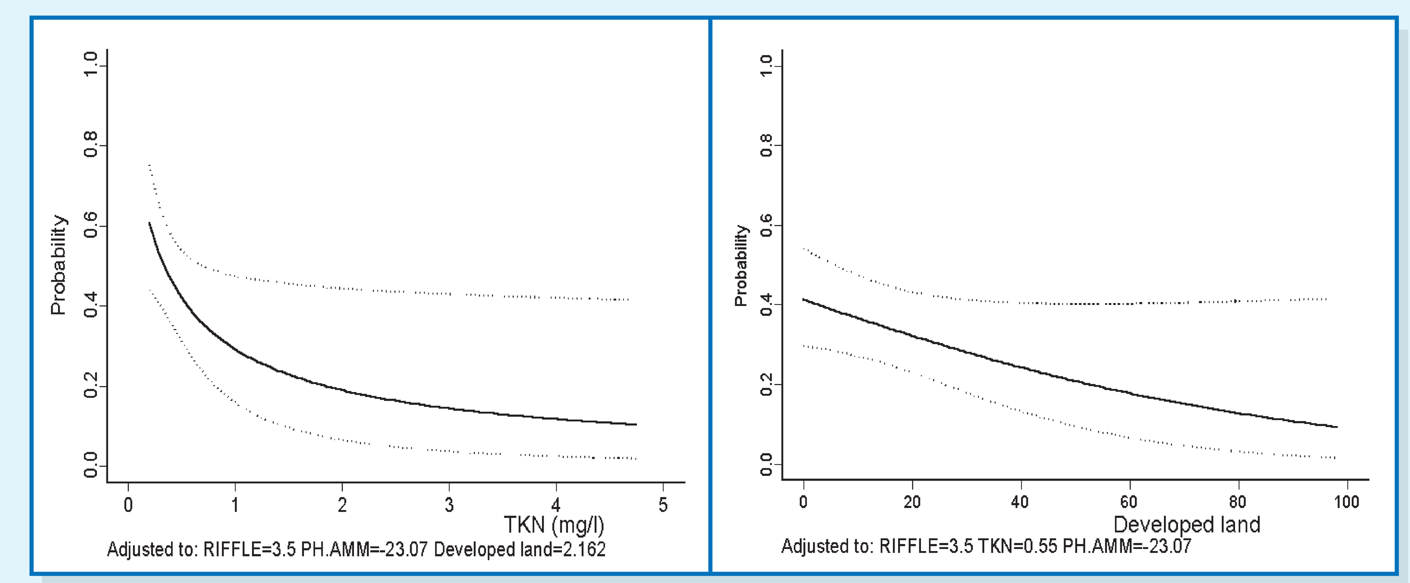
$\log(p2/(1-p2)) = -23.71 + 2.8\text{PH} - 0.70\log(\text{TKN}) + 0.03(\text{Forest})$

Total Chi square = 156, df = 9
Forward stepwise variable selection
Predictive efficiency = 60% compared to 25% using random assignment

Probability of ICI score in the "fair" category as compared to "poor"



Probability of Number Mayfly score being in "excellent" category compared to all other categories



Conclusions

- The results can be used for the identification of sites with a high risk of water quality degradation and for formal or informal "threshold development" for stressors.
- Current models were developed with the goal of "best predictive efficiency" using stepwise variable selection. This approach can be modified to develop models to predict the impact of stressor variables after accounting for natural variation. However, this modification would not necessarily result in "best predictive efficiency".
- Currently, prediction on each stressor variable while holding the others constant to their medians yields acceptable trends. However in most cases uncertainty represented by the prediction confidence intervals increases rapidly toward extreme values of the stressor variables. A realistic scenario would be to examine the distribution of estimated probabilities by varying the input variables simultaneously.
- This approach illustrates how secondary data like STORET and LEAPS can be valuable aides in identifying and predicting impacts.

Prepared by CSC for ORD/NERL/Cincinnati - November 2001